

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 29, 2002, 11:38:47 ; Search time 34.66 Seconds
(without alignments)
2073.420 Million cell updates/sec

Title: US-08-485-355B-50
3374
Sequence: 1 MGDAGVASQRPHNRRGRNV.....GKIARVARRARRARRAARAN 647

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq_032802:*

- 1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3374	100.0	647	19	AAW34535 Helicoverpa armigera
2	3353	99.4	647	15	AAR49662 Sequence of Heliot
3	289	8.6	634	19	AAW26785 Nudaurelia beta-li
4	289	8.6	634	19	AAW34537 Nudaurelia beta vi
5	139.5	4.1	1231	20	AAV00219 Enterococcus faeca
6	139.5	4.1	1265	20	AAV00218 Enterococcus faeca
7	127	3.8	2016	22	ABB66424 Drosophila melanog
8	121.5	3.6	2344	22	AAU37120 Staphylococcus aur
9	120.5	3.6	889	20	AAV24913 Human ontherin. H
10	120.5	3.6	889	20	AAV21687 Cathelin-like poly
11	120.5	3.6	979	22	AAE11961 Streptococcus agal

12	120	3.6	932	20	AAV00217 Enterococcus faeca
13	120	3.6	969	20	AAV00216 Enterococcus faeca
14	119.5	3.5	907	6	AAV50073 Epstein-Barr virus
15	119.5	3.5	907	21	AAV68009 Epstein-Barr virus
16	119.5	3.5	1425	22	ABW71908 Drosophila melanog
17	117.5	3.5	26926	22	AAU05396 Human titlin (cone
18	116	3.4	641	22	AAE00815 Aspergillus oryzae
19	116	3.4	1004	21	AAW23857 Hemophilus influenza
20	115.5	3.4	907	16	AAW80144 EBV gp350/220. Ep
21	113.5	3.4	658	21	AAV68010 Epstein-Barr virus
22	113.5	3.4	781	16	AAW08986 Human parvovirus V
23	113.5	3.4	816	12	AAW14444 A.faecalis penicil
24	112.5	3.3	503	22	AAW92291 C.glutamicum prote
25	112	3.3	889	22	AAW23879 Novel human diagno
26	111.5	3.3	583	22	AAW63470 Propionibacterium
27	110.5	3.3	1300	22	AAW33407 Enterococcus faeca
28	110.5	3.3	1300	22	AAU35313 Enterococcus faeca
29	110	3.3	703	21	AAW14781 Aspergillus oryzae
30	109	3.2	703	21	AAW14782 Aspergillus oryzae
31	109	3.2	780	20	AAV35502 Amino acid sequenc
32	109	3.2	1302	16	AAW84741 Multidrug resistan
33	109	3.2	1560	10	AAW94145 S.cremoris proteol
34	108	3.2	591	20	AAV23337 Talatromyces emerso
35	107.5	3.2	1472	22	AAW59423 Human polypeptide
36	107.5	3.2	2058	22	AAW97070 Human polypeptide
37	107	3.2	550	17	AAW89501 Penicillin acylase
38	107	3.2	550	17	AAW89502 Penicillin acylase
39	107	3.2	550	17	AAW89504 Penicillin acylase
40	107	3.2	550	17	AAW89499 Penicillin acylase
41	107	3.2	550	17	AAW89500 Penicillin acylase
42	107	3.2	690	21	AAV50835 A.oryzae glutamin
43	107	3.2	690	21	AAV50844 A.oryzae glutamin
44	107	3.2	812	12	AAW10047 abaa gene of Asper
45	107	3.2	2037	22	ABW71928 Drosophila melanog

ALIGNMENTS

RESULT 1	
ID AAW34535	standard; Protein; 647 AA.
AC AAW34535;	
XX	
DT 08-JUN-1998	(first entry)
DE Helicoverpa armigera virus coat protein p71.	
XX	
KW Vaccine; coat protein; p71; insecticide; Ig-like domain.	
XX	
OS Helicoverpa armigera virus.	
XX	
FH key	
FT Domain	Location/Qualifiers
FT	281..414
XX	/label= Immunoglobulin-like domain
XX	
PN W09746666-A1.	
XX	
PD 11-DEC-1997.	
XX	
PF 02-JUN-1997;	97WO-AU00349.
XX	
PR 31-MAY-1996;	96AU-0000234.
XX	
PA (CSTR) COMMONWEALTH SCT & IND RES ORG.	
XX	
PI Gordon KH, Hanzlik TN;	
XX	
DR WPI: 1998-042175/04.	
DR N-PSDB; AAT99117.	
XX	
PT Modified small RNA viruses and virus-like particles - have altered	

PT or substituted Ig-like domains to modify host cell tropism, useful
 as insecticides and in medicinal applications

PS Disclosure; Figure 1; 41pp; English.

XX The sequence is that of coat protein p71, it has an Ig-like domain
 CC and can be used in the production of virus-like particles (VLP). The
 CC VLPs can be used in vaccines where the Ig-like domain has been altered
 CC so that the VLP presents a surface located antigen which is used to
 CC elicit an immune response in a host organism. They can also be used
 CC for controlling the proliferation of a pest insect and potentially as
 CC medicinal delivery agents for cancer treatment and gene therapy.

XX Sequence 647 AA;

Query Match 100.0%; Score 3374; DB 19; Length 647;

Best Local Similarity 100.0%; Pred. No. 4,4e-273;

Matches 647; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGDAGVASQRPNRRGTRNVASANTVTVNGRRNRQRRTRGQVSPDNTFAAQLAOSL 60
 DB 1 mgdagvasqrpnmrrgrtrnvrsantvtvngrrnrrtrgrtrgrvqspdnftaaqdlagsl 60
 QY 61 DANTVTTPANISSMPERRNNAKGIIDSDSICGWFKYLDPAQTESARAVGESKIPDG 120
 DB 61 dantvttpanissmperrnnakgididsdsigwfkylpdaqtesaravgeskipdg 120
 QY 121 LVKFSVDAEIREINECEPVTVTVSVPLDGRQMSLIFSFPMFRTAVVAANVANKEMSL 180
 DB 121 lvkfsvdaelirelyneecpvttvsvpldgrqmslifsfpmtfayvaanvankemsl 180
 QY 181 DYNVDLIEMLNLDNRVYVDSQWINEFTNDTTYVIRIRLPRYDVPDTEGLVRFVSD 240
 DB 181 dvnvdliemlnldnrvyvdsqwinftndtlyvirirlyrpydvpdpdteglvrfvdsd 240
 QY 241 YRLTYKAITCEANMPTLVDSGFVIGGOYALTPTSLPOYDSEAVALTLPFAPSSAAL 300
 DB 241 yrltykailceaanmptlvdsqfvi99yaltptslpdydseavaltltpfapssaaal 300
 QY 301 AFVWAGLPQGTAPAGTPAMEQASSGGYLTWRHNGTTFPAGSVSYVLPFGFALERYDPND 360
 DB 301 afvwaglpqgstapagtpameqassggyltwrhngtffpagsvsyvlpegfalerydpnd 360
 QY 361 GSWTDFASAGDTVYFROYAVDEVVYTNPNPAGGSAFTFTRAVPSSNATNTVFNTLTET 420
 DB 361 gswtdfasagdtvfyfroyavdevvtnpnpgggsapftftravpssnayntlvfntltet 420
 QY 421 RPSRRLELPMPPADFCQTVANNPKIEOSLKERLGCYLVHSHKMRNPFOLTPASSFGAV 480
 DB 421 rpsrrlelpmppadfcqtvannpkieosllkerlgylvshskmrnpvfgltpassfgav 480
 QY 481 SFNNPGYERTDLPDYTGIRDSFDQNNSTAVAHFRSLSHSCSIVTKTYQMEGVNTNTP 540
 DB 481 sfnnpgyertdlpdytgirdsfdqnnstavafrslshscsivtktyqmevntnntp 540
 QY 541 FCGPAAHAGLKNKEIILCLADDLATRLGCVPAIDNFAANSAFPAANLSSVLSSEATSSI 600
 DB 541 fcgpaahagllkneellcladdlatrlgcvpatdnfaaasaafaannlssvlsseatsi 600
 QY 601 IKSGYERAVGAAGSLAKLPGLLMSYVGKITAARVARARRARRARRAAN 647
 DB 601 iksgyeravgaagslklp9llmsyvgkitaarvararrarraraan 647

RESULT 2

AA049662 standard; protein; 647 AA.

AA049662;

12-SEP-1994 (first entry)

XX

DE Sequence of Heliothis armigera RNA 2 p17.

XX HasV; RNA 1; small RNA virus; p17.

XX Heliothis armigera stunt virus.

XX WO9404660-A.

XX 03-MAR-1994.

XX 13-AUG-1993; 93MO-AU00411.

XX 14-AUG-1992; 92AU-0004081.

XX 08-JUL-1993; 93US-0085372.

XX (CSIR) COMMONWEALTH SCI & IND RES ORG.

XX (PACI-) PACIFIC SEEDS PTY LTD.

XX Christian PD, Gordon KHU, Hanzlik TN;

XX WPI; 1994-083180/10.

XX N-P-SDB; AAQ58523.

XX Small RNA virus capable of infecting insect species, e.g.

XX Heliothis - and transgenic plants contg. viral nucleic acid, for

XX protection against insect pests

XX Disclosure; Figure 2; 183pp; English.

XX The inventors claim a virus comprising a genome hybridisable with

XX the nucleotide sequence of RNA 1 or RNA 2; pref. the sequences are

XX those given in Figs 1 and 2 of the specification. As isolated

XX protein or polypeptide prepn. of the proteins or polypeptides

XX derivable from the virus are also claimed.

XX H. armigera larvae were raised and viral RNA was extracted. The virus

XX RNAs were reverse transcribed into cDNA. Clone hr236 contains about

XX 88% or RNA 2. A major translation product of apparent mol. wt. 24,000

XX is obt'd. This protein is derived from a mol. wt. 17,000 reading

XX frame overlapping the slab of the capsid protein gene. The Mr 24,000

XX protein (referred to as p17) may have a function in modifying or

XX manipulating the growth characteristics or cell cycle of

XX HasV-infected cells.

XX Sequence 647 AA;

Query Match 99.4%; Score 3353; DB 15; Length 647;

Best Local Similarity 99.5%; Pred. No. 2.5e-271;

Matches 644; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MGDAGVASQRPNRRGTRNVASANTVTVNGRRNRQRRTRGQVSPDNTFAAQLAOSL 60
 DB 1 mgdagvasqrpnmrrgrtrnvrsantvtvngrrnrrtrgrtrgrvqspdnftaaqdlagsl 60
 QY 61 DANTVTTPANISSMPERRNNAKGIIDSDSICGWFKYLDPAQTESARAVGESKIPDG 120
 DB 61 dantvttpanissmperrnnakgididsdsigwfkylpdaqtesaravgeskipdg 120
 QY 121 LVKFSVDAEIREINECEPVTVTVSVPLDGRQMSLIFSFPMFRTAVVAANVANKEMSL 180
 DB 121 lvkfsvdaelirelyneecpvttvsvpldgrqmslifsfpmtfayvaanvankemsl 180
 QY 181 DYNVDLIEMLNLDNRVYVDSQWINEFTNDTTYVIRIRLPRYDVPDTEGLVRFVSD 240
 DB 181 dvnvdliemlnldnrvyvdsqwinftndtlyvirirlyrpydvpdpdteglvrfvdsd 240
 QY 241 YRLTYKAITCEANMPTLVDSGFVIGGOYALTPTSLPOYDSEAVALTLPFAPSSAAL 300
 DB 241 yrltykailceaanmptlvdsqfvi99yaltptslpdydseavaltltpfapssaaal 300
 QY 301 AFVWAGLPQGTAPAGTPAMEQASSGGYLTWRHNGTTFPAGSVSYVLPFGFALERYDPND 360
 DB 301 afvwaglpqgstapagtpameqassggyltwrhngtffpagsvsyvlpegfalerydpnd 360

```

OY 361 GSWTDPASAGDTVTTPROVAVDEVVTVNNPAGGSAPTFTVRVPPSNAYNTNTERNTLEET 420
    |||||||
DB 361 gswtdpasagdtvttrqvavdevvtnmpaggsapltftvrppsnayntntrntllec 420
OY 421 RPSRRLELMPADPGGTAVANPKIEOSLKEKLECYLVHSMKRNPFQLPASSFGAV 480
    |||||||
DB 421 rpsrrlelmpadpfgtgvannpkieglkeltlgyclvshkmrnpvqlppassfgav 480
OY 481 SFNNNGCYERTRLPDYTGTRDSFDQNMSTAVAHFRLSHSCSIYVTKTYGMEGYNNVNP 540
    |||||||
DB 481 sfnnngcyertldpdytgtrdsfdqnmstavafrlshscslvltktygwegvtnvnp 540
OY 541 FEGFAHAGLKKNEELICLADDLATRLTGVPATDNFAAVSAFAANMLSSVLSKSEATSSI 600
    |||||||
DB 541 fegfahagllkneelcladdlatrltgyvpatdnfaaavsfaanmlssvlskseatsi 600
OY 601 IKSVEGTAVGAAGSLAKLPGLMSVPCKIAARVRRARRARRAARAN 647
    |||||||
DB 601 iksvgetavgaagslaklpglmsvpckiaarvrrarrarraaran 647

RESULT 3
AAM26785
ID AAM26785 standard; Protein; 634 AA.
AC AAM26785;
XX
XX 22-JUN-1998 (first entry)
DE Nudaurelia beta-like virus capsid protein precursor.
XX
XX NBV, RNA virus; transgenic plant; insect resistance;
XX disease resistance; Nudaurelia cytherea capensis;
KM pine tree emperor moth; capsid protein; vector;
KM virus-like particle.
XX
XX Nudaurelia beta-like virus.
XX
XX AU9724669-A.
XX
XX 04-DEC-1997.
XX
XX 02-JUN-1997; 97AU-0024669.
XX
XX 31-MAY-1996; 96AU-0000233.
XX
XX (CSIR ) COMMONWEALTH SCT & IND RES ORG.
XX (UKRH -) UNITV RHODES.
XX
XX Gordon KH, Hanzlik TN, Hendry DA;
XX
XX MPI: 1998-052736/06.
XX N-PSDB: AAV04471.
XX
XX Nudaurelia beta virus nucleic acid - useful for producing
XX recombinant virus, insect-resistant transgenic plants, etc
XX
XX Example 1; Fig 1; 33pp; English.
XX
XX This polypeptide comprises a 70 kDa capsid protein precursor whose
XX amino acid sequence was deduced from an open reading frame identified
XX in the RNA genome (see AAV04471) of Nudaurelia beta-like virus (NBV).
XX The capsid protein precursor is cleaved at a N/G site into 60.5 and
XX 8 kDa capsid proteins. A claimed infectious recombinant insect virus
XX vector comprises an expressible nucleic acid molecule comprising a
XX nucleotide sequence corresponding to all or an infectious and/or
XX insecticidal portion of the genomic RNA of NBV. Also claimed is: a
XX transgenic plant resistant to insect attack that produces NBV such
XX that insects feeding on the plant are deleteriously affected; and a
XX virus-like particle (VLP) prepared from expression of a nucleic acid
XX molecule comprising a sequence encoding the capsid protein of NBV.
XX The invention provides methods for the control of insect pests (the

```

```

CC pine-tree emperor moth Nudaurelia cytherea capensis is mentioned).
XX
XX Sequence 634 AA.
SO
Query Match 8.6%; Score 289; DB 19; Length 634;
Beat Local Similarity 25.7%; Pred. No. 2,8e-15;
Matches 177; Conservative 72; Mismatches 266; Indels 174; Gaps 39;
OY 35 QRRRTGROVSPPDNTAAQ--DLAOSLDANTVTTPANISSK-----PE----- 76
    || : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 31 qrrrtgrekpepratraaqtltsltsgeagkslprstdyqpawpnpereprehgqr 90
OY 77 --FRMKAKKIDLD--SDSIGMYFKYLDPPACATGESARAVGYSXIRPGLVFSVDAEIR 131
    || : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 91 sdltregakasdgeshgsdikawhndyldpdeyxtslddg---kipdgalpqscgcqgr 147
OY 132 ETYNEECPVYTDVSVPLDGRQMSLISFSPMERTRVAVAVANENKMSLDVYN-DL---- 186
    || : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 148 gtvgartypplnstlplddgtwpllvmbhlpfrrhlffltsnte--vevtnadlda 205
OY 187 IEMLNNLADKRYVDSQGMINTNTTYVRIKRLRP--YDVPDPE-GLVKTVSDYR 242
    || : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 206 ndw-nmrdtweat-ypswagvgn--vfywv---ptealtdvpptqlgvsqllseyr 257
OY 243 LTYKAITCEANMPRTLVDOGFWIGGOYALPTSLPOYDSEAVALTHLT-----FARPS 295
    || : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 258 ltsqytlayfnapllvngvavlaqf--qpdkenhknpdlvaqtlcgqtlqigsqpn 315
OY 296 SAALAFVWAGLPQGGTA-PAGTPAMEQASSGQYLWRHNGTTPAG--SVSYVLPBGF 351
    || : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 316 ycltmt-lgdyefgsgaalplptvmgmpesqglvfqnaultdfgvntltltlppgs 374
OY 352 ALERDIPNDGSMTPDASAG-DVITERQVAVDEVVTVNNPAGGSAPTFTVRVPPSNAYTN 410
    || : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 375 v-----tgmwqfcaasngldtvt-----vd-----aga-----lvry----- 400
OY 411 TVFRNTL--LETRPSSRREL-----PMPPADGGQTVANPKIEOSLKEKLECYLVHSM 464
    || : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 401 ----rsefgrlgesaghnqngdstndmpndag-----naktlqfqltkr--ghympesi 451
OY 465 RNPVFQLPASSFGAVSFNN-----PQYER--TRD-LPDYTGIRDSFDQNM 507
    || : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 452 r---vrndnatatsygvdededncgrlhraigaigyhqrdrlpmnty-----m 500
OY 508 STAVAHFRSLSHSCSIYVTKTYGMEGV-----TVNPFQGFAGLKKNEELICLADD 561
    || : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 562 LATRLTGVPATDNFAAVSAFAANMLSSVLSKSEATSSIRKSVGETAVGAAGSLAKLP 621
    || : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 545 wdlhlpfayperyngfgalfamvaktiaqpr-----yhrs-----aagvanavdcles 594
OY 622 LMSVPGKTAARVRRARRR-----RAAR 645
    || : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 595 atesvasnstsergrarrvvggiargar 623

RESULT 4
AAM34537
ID AAM34537 standard; Protein; 634 AA.
XX
XX AAM34537;
XX
XX 08-JUN-1998 (first entry)
DE Nudaurelia beta virus coat protein p70.
XX
XX Vaccine: coat protein; p70; insecticide; Ig-like domain.
XX
XX Nudaurelia beta virus.
XX
XX key Location/Qualifiers

```

FT	Domain	285..433
XX		/note= "Immunoglobulin-like domain"
XX		
PN	W09746666-A1.	
PD	11-DEC-1997.	
XX		
PF	02-JUN-1997;	97WO-AU00349.
XX		
PR	31-MAY-1996;	96AU-0000234.
XX		
PA	(CSIR) COMMONWEALTH SCT & IND RES ORG.	
XX		
PI	Gordon KH, Hanzlik TN;	
XX		
DR	MP1: 1998-042175/04.	
XX	N-PSTDB; AAT99118.	
XX		
PT	Modified small RNA viruses and virus-like particles - have altered	
PT	or substituted Ig-like domains to modify host cell tropism, useful	
XX	as insecticides and in medicinal applications	
PS	Disclosure; Figure 2; 41pp; English.	
XX		
CC	The sequence is that of coat protein p70, it includes an Ig-like domain	
CC	which can be used in the production of virus-like particles (VLPs). The	
CC	VLPs can be used in vaccines where the Ig-like domain has been altered	
CC	so that the VLP presents a surface located antigen which is used to	
CC	elicit an immune response in a host organism. They can also be used	
CC	for controlling the proliferation of a pest insect and potentially as	
CC	medicinal delivery agents for cancer treatment and gene therapy.	
XX		
SQ	Sequence	634 AA:

Query Match	8.6%	Score 289;	DB 19;	Length 634;
Best Local Similarity	25.7%	Pred. No. 2.8e-15;		
Matches 177;	Conservative 72;	Mismatches 266;	Indels 174;	Gaps 39

```

0Y 0RRRTGROVSPPDNFTAAAO--DLAOSLDANVTPEANISSM-----PE-----76
Db 31 qradktreagpepratraacqtlttscegaagsaktisptrstdyaparwmprehpgqr 90
Qy 77 --FRWAKGIDLD--SDSIGWFYKYLDPAGATESARAVGEYSKIPDGLVKFSVDAEIR 131
Db 91 sdtregakssddgeshsdtkawlbndyldpdgeyxctlddg---kxpdgalpqstcgqfr 147
Qy 132 ELYNECCPVYTVSVPLDGROMSLSFSPMPRTAIVAVANNENKMMSLDVVA-DL-----186
Db 148 glvgarjypjlnstclpldgdvlpwllymbhlpfirhllfllttsnte--vevtadadafa 205
Qy 187 IEMLNLLADMWRVYNVDSQWINFTNDPTYYVRIRLRPT--RDVPORTE-GLAVRVSDDR 242
Db 206 ndw-nmrctwcaat-ypsvaqygn--vfymvv-----ptcalcdvpptqlgysgllesyr 257
Qy 243 LTYKAITCEANNPRLVDGFWIGGOVALPRTSLPOYDSEAYALATLT-----FAPRS 295
Db 258 ltsqvtaeyfnaplvlngvavlaeqf--qpdkenhqenpdlvagttcgtgylqigsogn 315
Qy 296 SAAALAFWMAIGPOGTA-PACTPAHEQAQSSCGYLTWRINGTTFPAG---SWSYVLPEGF 351
Db 316 yllent-lgdqvefgaaalprlvtvmgmppesqglfcftanllfdygnltitcltlpprs 374
Qy 352 ALERDPMDSJTDSPASAG-DTWFERQVAVDEBVTNNPAGGSAPTFRVVRPPSNAYTN 410
Db 375 v-----tgmqfctasngtdclvt-----vd-----aga-----lvrv-----400
Qy 411 TVFRNTL-LLETRPSSRLLET---PMPADFGQTAVNPNKITEQSILKETGCYLVHSM 464
Db 401 ---rsefgrlgtesghqngdsctndmnpdag---naktlqfqgltxr--ghympaeai 451
Qy 465 RNPVQLRPASSFGAVSFNN-----FEYER--TRD-LPDYTGTGRDSFDQM 507

```

```

Db      452  r--vrdnnaesypvdedtedednglrhlraigalqgyhrqglrdlpsmng-----m 5000
Oy      508  STAAVHPSLSHSCSIYTKTYQGEGV-----TNVNTPGQQAHAAGLAKNEETICLADD 561
Db      501  stswpyr-----lqgyplrlstdtgggeggpffasatprpddvaltwart 544
Oy      562  LATRLTGYTPATDNPAAVSAFAANMLSSVLKASRTSIIKSVETINVGAAQGLAKLPG 621
Db      545  wcdlhpfeayperayngfgallfemavaktlaqldpr-----ayvanavtdcles 594
Oy      622  LMSVPGKIAARVARRARR-----RAAR 645
Db      595  atesvaanststerrqrratrrayvgfilargar 623

```

RESULT	5
AAV00219	
ID	AAV00219 standard; Protein; 1231 AA
vv	

AC	AAV00219;
XX	
DT	20-APR-1999 (first entry)

DE	Enterococcus faecalis antigenic polypeptide fragment EF108.
XX	
RW	Enterococcus faecalis; infection; vaccine; immune response; diagnosis;
RW	detection; attenuation; antigenic.
XX	
OS	Enterococcus faecalis.

PN W09850554-A2

PD 12-NOV-1998

PF 04-MAY-1998; 98WO-US08959.

14-NOV-1997: 97JUS-0066009

PR	06-MAY-1997;	97US-0044031-
PR	16-MAY-1997:	97US-0046655

XX
XX
(UTM2 -) HUMAN GENOME SCT TNG

XX

XX

DR N-PSDB; AAX20209.

PT New isolated Enterococcus faecalis polynucleotides - used to develop

PT for prevention or attenuation of *Enterococcus* infection

XX
DE
Claim 8: Page 211: 301ms: Englisch

XX
XX

CC isolated from *Enterococcus faecalis*. The present invention describes

CC The proteins can be used in vaccines for preventing or attenuating an

They can also be used for detecting *Enterococcus* antibodies in a sample

CC acids products from the present invention can also be used for
CC the nucleotide sequences can be used for detecting *Enterococcus* nucleic

screening compounds to identify agonists and antagonists of E. faecalis

3 X X
2
1
1
1
2
2
2
2
2
2

Query Match 4.18; Score 139.5; DB 20; Length 1231;

Best Local Similarity 15.06, Freq. NO: 0.024,
Matches 141; Conservative 94; Mismatches 252; Indels 257; Gaps 31;

QY 19 NRVYSANTVTVNGRRNRQRRRTGKOVSPPDNFTAAQDLAOSLDANTV----TEPANISS- 73

```

Db 542 nrkventfdngak-----ltpcftgqkktvls-daytfkagaglpdtytg 591
QY 74 --MPEFRMNAKGIIDSDSIGWFKYLDIPACATSAARAVGSKYRIPDLVFSVDAETR 131
Db 552 gktyfkgykxkslntlt-----tkapsyqvdydnddn 629
QY 132 EYNEECPVNDVSVPLD-----GROMSLSFSPMERTAVAVANVEN 175
Db 630 vyeetvtytvsdmtfveknkgafpalatfsqkyagqsay--lrdtydvtkskn 687
QY 176 -----KEMSLDVNDLIEMLN----- 191
Db 688 gngqytsvinsngsmplsgellkkyngqplsatnrlqfnvdklaidqqlkyvdsiqlda 747
QY 192 ---NLADMRVYV-----DSEQWNIENFDNTYYV---RIRVLRLP 223
Db 748 gsnlksyryyytnmslsvfdpnvapaevdlsseelnlinfoqgcltysnannrlfytchl 807
QY 224 TYD-VPDTEGLVRTVSD-----YRLTYKAITCEANMPRLVDGFWIGGYALTP 272
Db 808 gysgtpgyvnyllvmflfnakpadkelykvtkqvt-----entvd---vngakltap 858
QY 273 TSLPOYDVSEAYALHTLTFARPSSAAL-----AFWAGLPQGGTAPGTPAMEGA 323
Db 859 tgfthg---gngvpmnsntf-kylaakalpatytlgkykyltfgwykgtklkpsltl----- 908
QY 324 SSGGYLTWRHNGTTPAGSVSYVLEPGFALERYDNDGSMWTFASAGDPTVFRQYAVDEV 383
Db 909 -----nktcptfnat-----fdgndamtamykeeiaptasvltlrpkev 947
QY 384 VVTNNPAGGSAPTFTVVRPSPNAYNTVFRNTLTLETRPS--SRRLRLP--MPPADGQTV 440
Db 948 idtnlnvltwtltntskap-----lqnltlkqpnweagltlptfmevrepgeett 998
QY 441 ANNPKIEOSLKEITGCTLVHNSKMNPPVQULTPASSFGAVSP-----NNPGIERRDLDP 495
Db 999 kslp-vnstlwltegy-----plpnavpiglikvsvaalttratzgkntvlkaevv 1046
QY 496 YGINDSPDQNMSTVAHFRSLSHSCSIYTKYOGMEGTVNTNTPFGOPRHAHLKNEEI 555
Db 1047 fggikdstvndf-----vrlpndqdevvptegtfisvptfd--fgqygvagtgkqghel 1098
QY 556 LCLAD-----DLATRLTGVPATDNF-----AAVASAFAN 586
Db 1099 kqaadygngutrnpylrlkktqpnwsltaqjsqksatdplatrclllgaapvasfny 1158
QY 587 MLSVLKSF--ATSSIIKSVGETA 608
Db 1159 ngpteikntvgltisailanla 1182

```

```

* RESULT 6
AA00218
ID AAY00218 standard; Protein: 1265 AA.

```

```

XX AC AAY00218;
XX DT 20-APR-1999 (first entry)
XX DE Enterococcus faecalis protein EF108.
XX KW Enterococcus faecalis; infection; vaccine; immune response; diagnosis;
XX KW detection; attenuation; antigenic.
XX OS Enterococcus faecalis.
XX PN W09850554-A2.
XX PD 12-NOV-1998.
XX PF 04-MAY-1998; 98MO-US08959.
XX PR 14-NOV-1997; 97US-0066009.

```

```

PR 06-MAY-1997; 97US-0044031.
PR 16-MAY-1997; 97US-0046655.
PA (HUMA-) HUMAN GENOME SCT INC.
PI Bailey C, Choi GH, Hromockyj A, Kunsch CA;
DR WPI: 1999-070095/06.
DR N-PSDB; AAX20208.
XX
XX New isolated Enterococcus faecalis polynucleotides - used to develop
XX products for the detection of Enterococcus and for use in vaccines
XX for prevention or attenuation of Enterococcus infection
XX
XX Claim 9; Page 209-210; 301pp; English.
XX
XX The present sequence represents a protein isolated from
XX Enterococcus faecalis. The present invention describes genes, proteins
XX and antigenic polypeptides isolated from E. faecalis. The proteins can
XX be used in vaccines for preventing or attenuating an infection caused
XX by a member of the Enterococcus genus in an animal. They can also be
XX used for detecting Enterococcus antibodies in a sample. The nucleotide
XX sequences can be used for detecting Enterococcus nucleic acids.
XX CC products from the present invention can also be used for screening
XX compounds to identify agonists and antagonists of E. faecalis protein
XX activity.
XX
XX Sequence 1265 AA;
SQ

```

```

Query Match 4.1%; Score 139.5; DB 20; Length 1265;
Best Local Similarity 19.0%; Pred. NO. 0.025;
Matches 141; Conservative 94; Mismatches 252; Indels 257; Gaps 31;

```

```

QY 19 NRVSAANTYTVNGRRNORRRTGRQVSPDNPFAAADLAQSIDANTV---TFPNIIS- 73
Db 574 nrkventfdngak-----ltpcftgqkktvls-daytfkagaglpdtytg 623
QY 74 --MPEFRMNAKGIIDSDSIGWFKYLDIPACATSAARAVGSKYRIPDLVFSVDAETR 131
Db 624 gktyfkgykxkslntlt-----tkapsyqvdydnddn 661
QY 132 EYNEECPVNDVSVPLD-----GROMSLSFSPMERTAVAVANVEN 175
Db 662 vyeetvtytvsdmtfveknkgafpalatfsqkyagqsay--lrdtydvtkskn 719
QY 176 -----KEMSLDVNDLIEMLN----- 191
Db 720 gngqytsvinsngsmplsgellkkyngqplsatnrlqfnvdklaidqqlkyvdsiqlda 779
QY 192 ---NLADMRVYV-----DSEQWNIENFDNTYYV---RIRVLRLP 223
Db 780 gsnlksyryyytnmslsvfdpnvapaevdlsseelnlinfoqgcltysnannrlfytchl 839
QY 224 TYD-VPDTEGLVRTVSD-----YRLTYKAITCEANMPRLVDGFWIGGYALTP 272
Db 840 gysgtpgyvnyllvmflfnakpadkelykvtkqvt-----entvd---vngakltap 890
QY 273 TSLPOYDVSEAYALHTLTFARPSSAAL-----AFWAGLPQGGTAPGTPAMEGA 323
Db 891 tgfthg---gngvpmnsntf-kylaakalpatytlgkykyltfgwykgtklkpsltl----- 940
QY 324 SSGGYLTWRHNGTTPAGSVSYVLEPGFALERYDNDGSMWTFASAGDPTVFRQYAVDEV 383
Db 941 -----nktcptfnat-----fdgndamtamykeeiaptasvltlrpkev 979
QY 384 VVTNNPAGGSAPTFTVVRPSPNAYNTVFRNTLTLETRPS--SRRLRLP--MPPADGQTV 440
Db 980 idtnlnvltwtltntskap-----lqnltlkqpnweagltlptfmevrepgeett 1030
QY 441 ANNPKIEOSLKEITGCTLVHNSKMNPPVQULTPASSFGAVSP-----NNPGIERRDLDP 495
Db 1031 kslp-vnstlwltegy-----plpnavpiglikvsvaalttratzgkntvlkaevv 1078

```

```

OY 496 YGIGRDSFGQNNSTAAVAHFRSLSHOSCIYTKYQGEQVTNVTPEPGOAHAGLITNEEI 555
Db 1079 fggikstvdnf-----vrlprnqdevvrlptegflfsvptfd--fgqyvgatckqhs! 1130
OY 556 LCLAD-----DLATRLTGVPATDNF-----AAVASFAAN 586
Db 1131 kgaadygngtrnpyrlrkktpnws!taqlsqpkasatds!ptrtrlllgaapvsfny 1190
OY 587 MLSVLKSE--ATSSITKSGVERA 608
Db 1191 ngptelkntvgtsalsalcnm!ta 1214

RESULT 7
ABB66424
ID ABB66424 standard; Protein; 2016 AA.
XX
XX ABB66424;
XX
XX DT 26-MAR-2002 (first entry)
XX
XX DE Drosophila melanogaster polypeptide SEQ ID NO 26064.
XX
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX
XX KW pharmaceutical.
XX
XX OS Drosophila melanogaster.
XX
XX PN WO200171042-A2.
XX
XX PD 27-SEP-2001.
XX
XX PF 23-MAR-2001; 2001WO-US09231.
XX
XX PR 23-MAR-2000; 2000US-191637P.
XX
XX PR 11-JUL-2000; 2000US-0614150.
XX
XX PA (PEKE ) PE CORP NY.
XX
XX PI Venter JC, Adams M, Li PMD, Myers EW;
XX
XX DR WPI: 2001-656860/75.
XX
XX DR N-PSDB; ABLI0527.
XX
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX
XX PT interactions -
XX
XX PS Disclosure: SEQ ID NO 26064; 21pp + Sequence Listing; English.
XX
XX CC The invention relates to an isolated nucleic acid detection reagent
XX
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX
XX CC useful in developmental biology and in elucidating cell signalling and
XX
XX CC cell-cell interactions in higher eukaryotes for the development of
XX
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX
XX CC discloses genomic DNA sequences (AB101840-AB116175), expressed DNA
XX
XX CC sequences (AB101840-AB116175) and the encoded proteins
XX
XX CC (AB57737-AB872072).
XX
XX CC The sequence data for this patent did not form part of the printed
XX
XX CC specification, but was obtained in electronic format directly from WIPO
XX
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 2016 AA;

Query Match 3.8%; Score 127; DB 22; Length 2016;
Best Local Similarity 20.7%; Pred. No. 0.57;
Matches 77; Conservative 45; Mismatches 116; Indels 134; Gaps 18;

231 TGGIVRTVSVDYRLTYKAIKTCANMPTLVDOGFMI-----GGQVALPTPTSLPQYVSEA 283
1235 tttkttqkvpnymsfdeateleknkp---yetlwtastltlgeqgskslvampsdqvpak 1314

```

QY	284	YV----	LHITLFARRSSAAALAFWAGLPGCGTAPACTPAMWEDASSGCVLTWRHNGTTP	333
Db	1315	lasfdcltlatcfkedakmpcla----	vgapq-----	peitwkikgvelfs 135
QY	340	AGSVSVLPEEGFAL-----	ERYDPNDGS-WTDFASAGDVTFRQVAV-----	380
Db	1355	andmrtrvlpdgslllkxvnrgdagdyoschaenslakdstlkhllvlpapsphvtlsatt	141	
QY	381	-DEVVYTVNNPAGGSAAPL--	FTVAKVPS-----	NAYTN 410
Db	1415	lcalvklkphnegdaphghylhkpgefeweelsevsvdsgkhnlegllcgarygyva-	147	
QY	411	TVFRR-----	TLLETPPSSRRLELMPMPADFCOTVANNPKIEOSLTKETLGC----	TL 459
Db	1474	tgfmimgaesdlnlntctkqkpkipekr-flevsnsvsjhfakwdg-gcpmshty	153	
QY	460	VHSKMRNP-----	VFQLPAPASSEG--AVSFNNPGY-----	487
Db	1533	veskkrrddjlemqgsnmvdkpdmnyvldlepactynlrltahnsgatvayefactlvt	159	
QY	488	----	ERTRDLPD 495	
Db	1592	gqtlaprsdlpe	1603	
RESULT	8			
ANU37120				
ID	ANU37120	standard; Protein: 2344 AA.		
XX	ANU37120;			
XX	DT	14-FEB-2002 (first entry)		
XX	DE	Staphylococcus aureus cellular proliferation protein #1290.		
XX	KW	Antisense: prokaryotic cellular proliferation protein;		
XX	KM	antibiotic; antibacterial; drug design.		
OS	Staphylococcus aureus.			
PN	WO200170955-A2.			
XX	PD	27-SEP-2001.		
XX	PF	21-MAR-2001; 2001WO-US09180.		
PR	PR	21-MAR-2000; 2000US-191078P.		
PR	PR	23-MAY-2000; 2000US-206848P.		
PR	PR	26-MAY-2000; 2000US-207727P.		
PR	PR	23-OCT-2000; 2000US-242578P.		
PR	PR	27-NOV-2000; 2000US-253625P.		
PR	PR	22-DEC-2000; 2000US-257931P.		
PR	PR	16-FEB-2001; 2001US-269308P.		
XX	PA	(ELIT-) ELITRA PHARM INC.		
XX	PI	Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ,		
PI	PI	Yamamoto RT, Xu HH;		
XX	XX	WPI: 2001-611495/70.		
DR	DR	N-PSDB; AAS54979.		
XX	PT	New polynucleotides for the identification and development of		
XX	PT	antibiotics, comprise sequences of antisense nucleic acids -		
XX	PS	Example 3; Seq ID No 12713; 51pp; English.		
XX	XX	The invention relates to antisense inhibitors of genes essential to		
CC	CC	prokaryotic cellular proliferation, their use in identifying the		
CC	CC	genes, their use in the discovery of novel antibiotics, the essential		
CC	CC	genes themselves and the encoded proteins. The prokaryotes used are		
CC	CC	Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella		


```

Db 157 hdpdagenglrtylltrddnglfgldvksrgdgtkfpelvikald---teqgnhhtlvi 213
Oy 142 TDVSVPLDGRQMSLSIFSPMEFRTAVVAANVENKEMS--LDVYNDLIEWNLNADRRYV 199
Db 214 t-----aldgse-----pprsatvglnvkvldsndspvfeapsylvelpentpigtv 262
Oy 200 VDSEQMINFTN-----DFTY-----YVRIRVRPRYVDPPREGIVRTVS--DY----- 241
Db 263 id-----lnatdagedpgrgevlvysfsyvpdrv-relfsl-dpkrqgltrvkgndlveengm 316
Oy 242 -----RUTYKAITCEANMPRLVDQGWIGQVYALPTSPQDVSE 282
Db 317 leldvgardlgnllphckvkvtkldrndnapsi-----vsvrgalse 363
Oy 283 AYALHTLTFARPSSAALAFVW---AG-----LPQGTAPAGTPAMEQASSGGYLT 330
Db 364 a-----apgtvlalvrvldrdsgknqqlgcrlvggggtg99g----- 401
Oy 331 WRHNGTTFPAGSVSYVLPESFALERDNDGSDTDFASAGDTYTRQV---ANDEVVYTN 387
Db 402 ----glgpgsgsvpfkleeny-----dnfyt-----vvtdrpmrdetqdeynvtl 442
Oy 388 NPAGGGSAP-----TFTVRY-----P-----SNAYTNTVFRNTLLETRSSRLLEPMP 433
Db 443 vardgsgprlnskstafaklildndnprftkglvlyqvhennl-----pg 488
Oy 434 ADFGQTFANPKIEQ-----SLKETLGCYLVHSHKMR--NPVEQLPPASSFGAVSFNMPG 486
Db 489 dsjgsvlaqgpdjgngltvgsyllpshlgtvsltytysvnp---tngalylalrfsfn--- 541
Oy 487 YERTRLDLYTGIRDS-----FDONMSTAVAHFRSLSHSCSYTKTYQGEG----- 533
Db 542 feqlkafefvklakdsapahlesnatlvrvtlvdndnapvlyprltlgnltaelqyvrna 601
Oy 534 -----VTNVNTPRPGOFANGLTK-----NEELICLADDLATRLTGYPATDN-----F 576
Db 602 gllylvsvtvaldsdfigesgrltyelvdgnddnlfeidpsgegrltnlprfmedvtprvel 661
Oy 577 AAAYSAFANMLSSVYLSEATSSIIKSVGETAVGAQASGLAKLP 621
Db 662 vvvtdhngkprltlsavak-----llfrsvs-----gslpegvpyrvg 697

RESULT 10
AAZ1687
ID AAY21687 standard: Protein; 889 AA.
XX
XX AAY21687;
XX
XX 18-AUG-1999 (first entry)
XX
XX Cadherin-like polypeptide, ontherin.
XX
XX Ontherin: cadherin-like polypeptide; cadherin: cell differentiation;
XX neuronal cell; testicular; renal; spermatogenesis; vertebrate; stroke;
XX nervous system; neurological; injury; ischemia; inflammatory; tumour;
XX Alzheimer's disease; neurodegenerative disease; Parkinson's disease;
XX Huntington's chorea; amyotrophic lateral sclerosis; multiple sclerosis;
XX spinocerebellar degeneration; pain syndrome; drug screening.
XX
XX Vertebrata.
XX
XX WO9929853-A1.
XX
XX 17-JUN-1999.
XX
XX
XX 08-DEC-1998; 98WO-US25981.
XX
XX 08-DEC-1997; 97US-0067887.
XX
XX (GENY ) GENETICS INST INC.
XX

```

```

PI Israel DI;
XX
XX WPI; 1999-385603/32.
DR N-PSDB; AAX77400.
XX
XX New isolated cadherin-like polypeptides useful for treating
XX Alzheimer's disease
XX
XX Claim 20: Page 96-101; 108pp; English.
XX
XX This represents a cadherin-like polypeptide, ontherin, which can bind
XX to at least one of Ca2+, a catenin or a cadherin. The ontherin (OT)
XX protein regulate differentiation of neuronal cells; regulate survival of
XX differentiated neuronal cells; regulate proliferation of testicular germ
XX line cells; and/or regulate proliferation of renal cells. The polypeptide
XX preferably regulate spermatogenesis. The OT polypeptides are involved in
XX the formation and maintenance of ordered spatial arrangements of
XX differentiated tissues in vertebrates, both adult and embryonic, and can
XX be used to generate and/or maintain an array of different vertebrate
XX tissue both in vitro and in vivo. OT therapeutics can be used for
XX treating e.g. neurological conditions deriving from acute, subacute, or
XX chronic injury to the nervous system, including traumatic, chemical, and
XX vascular injury and deficits (such as the ischemia resulting from stroke).
XX together with infectious/inflammatory and tumour induced injury; aging of
XX the nervous system such as Alzheimer's disease; chronic neurodegenerative
XX diseases of the nervous system (Parkinson's disease, Huntington's chorea,
XX amyotrophic lateral sclerosis as well as spinocerebellar degenerations);
XX and chronic immunological diseases of the nervous system or affecting the
XX nervous system including multiple sclerosis, for selective ablation of
XX sensory neurons, e.g. in the treatment of chronic pain syndromes, or in
XX the treatment of neoplastic or hyperplastic transformations such as may
XX occur in the central nervous system. The products may also be used in
XX other therapeutic applications related to their activities. The products
XX can also be used for detection, diagnosis and drug screening.
XX
XX Sequence 889 AA:
XX
XX Query Match 3.6%; Score 120.5; DB 20; Length 889;
XX Best Local Similarity 20.4%; Pred. No. 0.57;
XX Matches 144; Conservative 92; Mismatches 250; Indels 219; Gaps 36;
XX
XX 32 RNRQRRRTGRQVSPPNFTFAADLQASLDANVTTPANISSPERFNNAKG----- 83
XX 97 rhnakcqlslevfandkeicmikevlqldndnapsltsdqdididseenaapgrtfrfltsa 156
XX 84 -KIDLDSDSIGWYFKYLDPADGATE-SARAVGEYSKIPDGKFSVDAIREIYNEGCPV 141
XX 157 hdpdagenglrtylltrddnglfgldvksrgdgtkfpelvikald---teqgnhhtlvi 213
XX 142 TDVSVPLDGRQMSLSIFSPMEFRTAVVAANVENKEMS--LDVYNDLIEWNLNADRRYV 199
XX 214 t-----aldgse-----pprsatvglnvkvldsndspvfeapsylvelpentpigtv 262
XX 200 VDSEQMINFTN-----DFTY-----YVRIRVRPRYVDPPREGIVRTVS--DY----- 241
XX 263 id-----lnatdagedpgrgevlvysfsyvpdrv-relfsl-dpkrqgltrvkgndlveengm 316
XX 242 -----RUTYKAITCEANMPRLVDQGWIGQVYALPTSPQDVSE 282
XX 317 leldvgardlgnllphckvkvtkldrndnapsi-----gf-----vsvrgalse 363
XX 283 AYALHTLTFARPSSAALAFVW---AG-----LPQGTAPAGTPAMEQASSGGYLT 330
XX 364 a-----apgtvlalvrvldrdsgknqqlgcrlvggggtg99g----- 401
XX 331 WRHNGTTFPAGSVSYVLPESFALERDNDGSDTDFASAGDTYTRQV---ANDEVVYTN 387
XX 402 ----glgpgsgsvpfkleeny-----dnfyt-----vvtdrpmrdetqdeynvtl 442
XX 388 NPAGGGSAP-----TFTVRY-----P-----SNAYTNTVFRNTLLETRSSRLLEPMP 433
XX 443 vardgsgprlnskstafaklildndnprftkglvlyqvhennl-----pg 488

```


434 ADFQGTAVNNKRIED-----SLKETEGICGLVHSMKR-NPFPOLTPASSFGAVSPNNNG 486
 Db 489 dslgslvslqdpdlqgnqlgtvslslpshlqgdvsllyevsvnp-----lmgalylrshfn 541
 QY 487 YERTDLDLDYDYGIRPS-----FDQNMSTAVAFHFRSLSHSCSLVTKTYQWEG----- 533
 Db 542 feqtkafeafkylakdsqgphahlesnatrvrtyldvndnapevilylptlqndtelqyprna 601
 QY 534 -----VTNVNPFEGFPAHAGLKR-----NEELCLADDLATRLTGVPATDN-----F 576
 Db 602 glqylvsvtrvraldsqfsgsgrltyelvdgnddhlfeidpsgeirtlhpfwedvcpvel 661
 QY 577 AAAYSAFAANMSSVLKSEATSSIIKSGVETVGAAGSLALPG 621
 Db 662 vkvxvdhkgkpltsavak-----lltrsvs---gsllpegyprvng 697
 RESULT 11
 AAEL1961
 ID AAEL1961 standard; Protein: 979 AA.
 AC AAEL1961;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Streptococcus agalactiae R5 protein.
 XX
 KW Cell surface protective antigen; R5 protein; antibacterial; therapeutic;
 XX immunotherapy; group B streptococci infection; prophylactic; vaccine.
 OS Streptococcus agalactiae.
 XX
 FN MO200173037-A2.
 XX
 PD 04-OCT-2001.
 XX
 PF 29-MAR-2001; 2001WO-EP03618.
 XX
 PR 30-MAR-2000; 2000EP-0106818.
 PR 17-OCT-2000; 2000EP-0122620.
 XX
 PA (GBFB) GES BIOTECHNOLOGISCHE FORSCHUNG MBH.
 XX
 PI Chhatwal GS, Erdogan S, Fagan PK, Guzman CA, Talay SR;
 XX
 DR WPI: 2001-616506/71.
 DR N-PSDB; AADI17326.
 XX
 PT Vaccine for preventing and treating infection by group B streptococci
 PT comprises the cell surface protective antigen of group B streptococci,
 PT R5 protein -
 XX
 PS Claim 11; Fig 1; 48pp; English.
 CC The invention relates to new cell surface protective antigen of group B
 CC streptococci, R5 protein and its corresponding DNA molecule. R5 protein
 CC and its DNA are useful for diagnosing an infection with group B
 CC streptococci. A composition comprising R5 protein is useful for
 CC immunotherapy of an infection with group B streptococci and for
 CC prophylactic and therapeutic vaccination against the same infection.
 CC The present sequence is Streptococcus agalactiae R5 protein.
 CC
 XX
 Sequence 979 AA;
 50
 Query Match 3.6%; Score 120.5; DB 22; Length 979;
 Best Local Similarity 22.2%; Pred. No. 0.66;
 Matches 107; Conservative 71; Mismatches 155; Indels 149; Gaps 30.
 19 NVKRS-AATVYVINGRRNRKRG-----QVSPDNFPAADLAQSIDANTV--- 65
 258 narlsftctvmengrlvekfdyarqsyntstffqdltpslstekassgvallvtantvdvt 317

```

0Y 66 ---TFPANISSMPREFRNNAKCKIDLDSDSIGWTFKYLDPRGATGESARANGESKIPDGLV 122
Db 318 dknitnesvvn-padedftvngp-drtpdav-----pag-tygsairtkgftc----- 359
0Y 123 KFSVDALFRELYNEECPVAVTVDSVPLDRQWKSLSIFSEPMFENFA---YVAVANVENKEMS 179
Db 360 ----daatk-lyhnavp-----staficekyhgdgyunsianvrtvka 399
0Y 180 LDVNDLLEWNLNLADMRYVDSBOWINFNDTTYVVRIRIVLRPTVDVDPDPTGELVRTVS 239
Db 400 pslngjlnkhasvdidyk-lsdnm---disnddlrlstrvlqprgsv-----vv 445
0Y 240 DYRLUTKAITCEANMPRTLVDDGFMIGGOMALTPRSL--OYDSEANA-LHT-LTFARPS 295
Db 446 ny-----ldcegnl-----lgyeyddttdairgthynlaessgdlinsdatverps 490
0Y 296 SAALALFWMALDPGGTAPACTPAPMEQASSGGLYTMHNGSTFEPAG-----S 342
Db 491 titkgkgkuydlvaeniltvpg-----kvnsdg-rlatngssfnugydaasaevaegtkS 543
0Y 343 VSYVL-----PEGALEREYDNDOSWIDFASAGDTYFROVANDDEVVYTNMNA----- 390
Db 544 vlyvyslikgeskgnvuhary-vilgtelelasa-ktvk-seapideaysckapartlekdgk 600
0Y 391 -----GGGSAP-----FTF-----VRVPPSNATNTVNFRTLLLETPSSRRLE 428
Db 601 lyefvhrvndngdadaadgkvtevegqtlitlyeyvervkrvgnvdyvlegtt-atrkpdtvyd 658
0Y 429 LP 430
Db 659 tp 660

```

RESULT	12
AAV00217	
ID	AAV00217 standard; Protein; 932 AA.
XX	
AC	AAV00217;
XX	
DT	20-APR-1999 (first entry)
XX	
DE	Enterococcus faecalis antigenic polypeptide fragment EF107.
XX	
KW	Enterococcus faecalis; infection; vaccine; immune response; diagnosis;
XX	detection; attenuation; antigenic.
OS	Enterococcus faecalis.
XX	
PN	WC9850554-A2.
XX	
PD	12-NOV-1998.
XX	
PF	04-MAY-1998; 98WO-US08959.
XX	
PR	14-NOV-1997; 97US-0066009.
XX	
PR	06-MAY-1997; 97US-0044031.
XX	
PR	16-MAY-1997; 97US-0046655.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Bailey C, Choi GH, Hromockyj A, Kunsch CA;
XX	
DR	WPI: 1999-070095/06.
XX	
DR	N-PSDB; AAX20207.
XX	
XX	
PT	New isolated Enterococcus faecalis polynucleotides - used to develop
XX	products for the detection of Enterococcus and for use in vaccines
PT	for prevention or attenuation of Enterococcus infection
XX	
PS	Claim 9; Page 208; 301pp; English.
XX	
CC	The present sequence represents an antigenic polypeptide fragment

CC isolated from *Enterococcus faecalis*. The present invention describes
 CC genes, proteins and antigenic polypeptides isolated from *E. faecalis*.
 CC The proteins can be used in vaccines for preventing or attenuating an
 CC infection caused by a member of the *Enterococcus* genus in an animal.
 CC They can also be used for detecting *Enterococcus* antibodies in a sample.
 CC The nucleotide sequences can be used for detecting *Enterococcus* nucleic
 CC acids. Products from the present invention can also be used for
 CC screening compounds to identify agonists and antagonists of *E. faecalis*
 CC protein activity.

XX
 CC
 XX
 SQ Sequence 932 AA;

Query Match 3.6%; Score 120; DB 20; Length 932;

Best Local Similarity 21.0%; Pred. No. 0.67;
 Matches 159; Conservative 82; Mismatches 261; Indels 254; Gaps 45;

QY 2 GDAGVASGRPHNRGTRNVRVSANTVYNGRRNORRT-----GRQVSPDNFTAAQ 54
 DB 222 gdvg-gptshqrrgtag-----pvyyltkrrvtekfempaggaipapeytc---q 269
 QY 55 DLAQSLDANTVTPANISME-----FRNNAKGI---DLSDSIGWY----- 95
 DB 270 dkktivgedfttqge-gtlperytsdgktylfkgykgnakpsletkttspsyavtyd 328
 QY 96 -----FKYLDPA-----AT--ESARAVGEYSKI 117
 DB 329 dnddlhvyeaavmkytlparealfgyvdeggnlmpakfklsatmgsgdgcgtemttf 388
 QY 118 P--DG-----LYKFSVDAIREIYN--ECCPVYT---DVSVPDLGRQMSISFSFPM 162
 DB 389 pldgldmpasqllkllaip--qkvycrpdgctivtvgpgevselpkyqyqltislsp-- 442
 QY 163 FRTAYVA-----VAN-----VENKEMSLDVNDLIEW---LNNLADMRVYVDSQWI- 206
 DB 443 -ttaytgdkckypvnevrriengnd--nlvsalvgxxaynlqksatryltarisywxw 498
 QY 207 -----NF---TNDTTYVRIRVLRPTVPP-----TEGL 234
 DB 499 gptcltysmslysgtaagnlystpdqcltlyyllenrrvtehfvdegakltptcftqgn 558
 QY 235 VRTSDVRLRY---KAI--TCEANMPTLVDOGFMIGGY-----ALRPTSLPOY 278
 DB 559 qlvvdseenyvvtvakalpklygaektylifgq-wfkgtktpalktcttspfcplfnded 617
 QY 279 DVSEAY-----ALHTLTPARP--SSAALAFAVWAGLPQGGTAPAPWADQASSGGYLT 330
 DB 618 dmavtgealptaeltltgavdlengatmdyweallkngeaplttkikltpa-----t 672
 QY 331 WRHNGTTPPAGSVSYVLPFGALERTYDPNDGSWTDFASAGTVPFRQYAVDEVVVTNMPA 390
 DB 673 w--aaglgapnltlfvgttgnlkafpvckeqwc--tgaqvsiltldp1-----pa 718
 QY 391 GG-----GSAPT-----FTVVRPPSNATNTVFRNTLTETRPSSRLLEP----- 430
 DB 719 gggqklmlltgvatgpnpgqvtladvevtgnfgstaktcd-vrkldlqelispdgdfis 777
 QY 431 MPADFGQTVANNPKEIOSLKETLGCYLVHSMKRNPFOL--TPASSGAVSFNNPGEYR 489
 DB 778 tptdfgklaisgsk-qgyllkkaady--gngrtrnpylrlntsganwslaqisqkxa 834
 QY 490 TRLDPTVTGIRDSFDQNMSTNAVHFRSLSHSCSLVTYTCQEGCVTVNVTNPFQGFANAGL 549
 DB 835 tdsiptrtlll-----lgtaaa-----asftdyngptctrcp1g----- 868
 QY 550 LKNEELICLADDLAT-----RLTG--VYPATONFA 577
 DB 869 -ktsvtlctadnctatavvanqfsgsdvylidftla 903

RESULT 13
 AAY00216
 ID AAY00216 standard; Protein; 969 AA.

XX
 AC AAY00216;
 XX
 DT 20-APR-1999 (first entry)
 XX
 DE *Enterococcus faecalis* protein EPI107.
 XX
 KW *Enterococcus faecalis*; infection; vaccine; immune response; diagnosis;
 XX detection; attenuation; antigenic.
 XX
 OS *Enterococcus faecalis*.
 XX
 PN W09850554-A2.
 XX
 PD 12-NOV-1998.
 XX
 PE 04-MAY-1998; 98MO-US08959.
 XX
 PR 14-NOV-1997; 97US-0066009.
 PR 06-MAY-1997; 97US-0044031.
 PR 16-MAY-1997; 97US-0046655.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Bailey C, Choi GH, Hromockyj A, Kunsch CA;
 PI WPI; 1999-070095/06.
 DR N-PSDB; AAX20206.
 XX

PT New isolated *Enterococcus faecalis* polynucleotides - used to develop
 PT products for the detection of *Enterococcus* and for use in vaccines
 PT for prevention or attenuation of *Enterococcus* infection
 PS Claim 9; Page 206-207; 301pp; English.

XX The present sequence represents a protein isolated from
 CC *Enterococcus faecalis*. The present invention describes genes, proteins
 CC and antigenic polypeptides isolated from *E. faecalis*. The proteins can
 CC be used in vaccines for preventing or attenuating an infection caused
 CC by a member of the *Enterococcus* genus in an animal. They can also be
 CC used for detecting *Enterococcus* antibodies in a sample. The nucleotide
 CC sequences can be used for detecting *Enterococcus* nucleic acids.
 CC Products from the present invention can also be used for screening
 CC compounds to identify agonists and antagonists of *E. faecalis* protein
 CC activity.
 XX
 XX

SQ Sequence 969 AA;

Query Match 3.6%; Score 120; DB 20; Length 969;

Best Local Similarity 21.0%; Pred. No. 0.71;
 Matches 159; Conservative 82; Mismatches 261; Indels 254; Gaps 45;

QY 2 GDAGVASGRPHNRGTRNVRVSANTVYNGRRNORRT-----GRQVSPDNFTAAQ 54
 DB 259 gdvg-gptshqrrgtag-----pvyyltkrrvtekfempaggaipapeytc---q 306
 QY 55 DLAQSLDANTVTPANISME-----FRNNAKGI---DLSDSIGWY----- 95
 DB 307 dkktivgedfttqge-gtlperytsdgktylfkgykgnakpsletkttspsyavtyd 365
 QY 96 -----FKYLDPA-----AT--ESARAVGEYSKI 117
 DB 366 dnddlhvyeaavmkytlparealfgyvdeggnlmpakfklsatmgsgdgcgtemttf 425
 QY 118 P--DG-----LYKFSVDAIREIYN--ECCPVYT---DVSVPDLGRQMSISFSFPM 162
 DB 426 pldgldmpasqllkllaip--qkvycrpdgctivtvgpgevselpkyqyqltislsp-- 479
 QY 163 FRTAYVA-----VAN-----VENKEMSLDVNDLIEW---LNNLADMRVYVDSQWI- 206
 DB 480 -ttaytgdkckypvnevrriengnd--nlvsalvgxxaynlqksatryltarisywxw 535

XX W09964603-A2.
 XX 16-DEC-1999.
 XX 10-JUN-1999; 99WO-US13113.
 XX 12-JUN-1998; 98US-0089158.
 XX (JACK-) JACKSON FOUND ADVANCEMENT MILITARY MED.
 XX (MOND/) MOND J J.
 XX (LEES/) LEES A.
 XX Mond JJ, Lees A;
 XX WPI: 2000-116546/10.
 XX Adjuvant containing an Epstein-Barr virus sequence that binds to
 PT complement receptor 2, for protecting against, e.g. infection or tumors
 PT
 XX
 XX Disclosure: Fig 1A: 43pp; English.
 XX
 XX The present invention describes an adjuvant (A) comprising at least one
 CC EBV (Epstein-Barr virus) gp350/220 sequence (I) that binds to CR2
 CC (complement receptor 2; CD21). Vaccines containing (A) and at least one
 CC other antigenic epitope (also similar immunogens and nucleic acid
 CC vectors that express (I)) are used to elicit an antibody response to
 CC both (I) (for protection against EBV) and the second epitope, which may
 CC be associated with some other infectious agent, allergen, tumour antigen
 CC or generally any condition that requires immune stimulation. The vaccines
 CC and immunogens may also be used to treat B cells in vitro. (A) increases
 CC the immunogenicity of a co-administered antigen. By combining (I), a
 CC non-specific B cell stimulator, with another antigen, the stimulatory
 CC effect is targeted to antigen-specific B cells, i.e. B cells that
 CC secrete immunoglobulins specific for the antigen are stimulated
 CC simultaneously through the antigen receptor and through (I)-mediated
 CC crosslinking of CR2. Constructs of (I) and second antigen are relatively
 CC small, so should be cleared relatively quickly (reducing immunogenicity)
 CC and many copies of second antigen can be included to increase
 CC antigenicity. The need for expensive anti-CR antibodies and potentially
 CC inflammatory C3 components is avoided and folding of the CR2-binding
 CC domain in (I)-antigen fusion proteins is unlikely to be disrupted. The
 CC present sequence represents EBV major outer envelope glycoprotein gp350.
 CC
 XX Sequence 907 AA:

Query Match 3.5%; Score 119.5; DB 21; Length 907;
 Best Local Similarity 19.5%; Pred. No. 0.71;
 Matches 136; Conservative 67; Mismatches 274; Indels 219; Gaps 27;

QY 48 NETAADODLAOSLDANTVTTPANISSMPEFRNMAKGIIDLSIGWFKYLDPAQATES 107
 DB 169 nltaavr--agld---vllplslpsaqsdenfs---vktmlyg---neid---lec 211
 QY 108 ARAVGEYSKIPDGLVKSVDARE-----IYNECPVVTDSVPLDGRQMSLISFSF 160
 DB 212 lmedgelsqylpgdnklnltcsqyeshpsglltsqvat--pdpqlyayslrltpr 269
 QY 161 PMFR-----TAYVAVANVENKESLD-----VYNDLI-----EMLNLADMRVYVDSSE 203
 DB 270 pvarflgnamlllyfygngpkasgdyclqsnlvfisdelpasqdmptntdltlyvgn- 328
 QY 204 QWINTNDTTYVYRIRVLPDYDPTTEGLVRTVSDYRLTYKAITCEANMPTLVDOGFV 263
 DB 329 -----atysvmvtse-----danspvtvtafw 352
 QY 264 -----IGQYALTP----- 273
 DB 353 awpntetdfckwltsgtspscenlsqafasnrtdftlvsgjgtapklilrtatna 412
 QY 274 -----SLPOVDSEAVALHTLTFARPSAAALAFVWAGLPQGTARAGTPAMQAS 324

DB 413 ttthkylfskapesetstspclntgtjadpnt-----tqlpssthyplnula--pas 463
 QY 325 SGGYLTMWNGTTPPAGSVSVLPEGFALERXDNDGSM-----TDFASADQTV 373
 DB 464 tqrvtstadvstspcpagtsqasp-----vcpsspwdngtleskapdmststspvtcp 516
 QY 374 TFRQVADEVVVT-----NNPAGGGSAPRPTTVVPPSNAYTNTVFR 414
 DB 517 tpnatspavatltptnatstptpavtptcpnatstptlgtksptstpsavtptcpnatstplgk 576
 QY 415 N-----TLETSPSSRRLLEMPADFGQTVANNPKIEOSLKETIGCY 458
 DB 577 tsptsavtltptnatstplgtkspts--avtltptnat--gplvgetsp--qanatnhtlg-- 631
 QY 459 LVHSMKMRPVFQLPASSFGAVSFNNNGYERTRLDPYTGIRDSFDQNMSTAVAHFRSL 518
 DB 632 ---gstptpvtstspkntatsavtctg--hntststsmstirpsnpeclstpsstsdnts 686
 QY 519 HSCSIWTKTYOGWEGVTNNTPFQGFAGLKNKEIILCLADLATRLTGVPATDNFAA 578
 DB 687 hmlptstahptggenltqvt--tpasistch-----hvtstsppeprptstgasqsgpnst 728
 QY 579 AVSAFAANMLSSVLKSEATSSIIISVGETAVGAQAS 614
 DB 739 slkpgvwnvltkypgnatstpqapsqktavptvts 774

Search completed: May 29, 2002, 11:39:37
 Job time: 50 sec